Vol. 13 No. 3 天 文 研 究 与 技 术 第 13 卷 第 3 期
Jul., 2016 ASTRONOMICAL RESEARCH AND TECHNOLOGY 2016 年 7 月

CN 53-1189/P ISSN 1672-7673

一种新的数据融合方法——广义融合法*

魏彦飞[1,2]，耿建平[1]，施浒立[2,3]，李景景[4]，李 林[2]，徐希坤[1]

（1. 桂林电子科技大学电子工程与自动化学院，广西 桂林 541004；2. 中国科学院国家天文台，北京 100012；
3. 北京日月九天科技有限公司，北京 100012；4. 中原工学院电子信息学院，河南 郑州 450007）

摘要：数据融合的目的是合并或者综合多个数据集，成为一个完整的性能更好的数据集。基于广义延拓插值外推法提出建立两组或两组以上方程解集和数据序列融合优化的数学模型，称之为广义融合法，此方法能充分发挥数据和量测资源的潜力，运用创新的集成和互校算法，解决数据融合面临的难题。介绍了广义融合数学模型的建立和求解方法的主要步骤；对模型中权值的分配原则做了分析；最后应用于卫星导航数据处理领域。实验证明该方法实用性好，数据平稳性好，外推稳定度高，有望成为一种新的解决数据融合问题的普适求解方法。

关键词：数据融合；广义融合；广义延拓外推；模型求解；权值分配；卫星导航数据处理

中图分类号：TN96 **文献标识码**：A **文章编号**：1672-7673(2016)03-0318-08

数据融合(也叫信息融合)的概念始于 20 世纪 70 年代，至 80 年代其技术取得了一定的进步和发展，到 90 年代不少数据融合技术的研究成果已得到实际应用，取得了理想的效果[1]。数据融合的主要原理是用多个传感器(同质或异质的)对同一对象进行描述，得到该对象的多源观测数据，将这些数据进行集成，以获得对被控对象的更好的理解[2]。

多传感器信息融合技术是近几年发展起来的一门实践应用技术，是针对一个系统使用多种传感器这一特定问题而展开的一种关于数据处理的技术[3]。对于多传感器系统，由于信息的多样化和复杂性，因此，信息融合方法的基本要求是具有强大的处理能力，以及较高的计算速度和精度。人工智能的发展使信息融合技术走向智能化、集成化。

数据融合的目的是合并或合成多个数据集，成为一个完整的性能更好的数据集。数据融合技术的实现，关键在于融合算法[4-5]。在融合求解过程中，不是仅仅把数据或解集简单地直接组合，这样做效果不会很明显，甚至最后得到的结果更差。为了能够得到良好的融合效果，应该分析各组方程解或多组数据集序列的特征。比如一些数据序列的随机误差较小但偏差较大；一些数据的随机误差较大但偏差较小[6-7]。这样就可以将随机误差较小的数据序列的相对变化量通过一定的方法融合到随机误差较大的数据序列中，能获得随机误差较小且偏差也较小的数据序列。本文基于广义延拓插值外推法[8]提出了一种求解一类特定的关联群方程组的多个离散数据集或最优融合估计集的方法，称之为广义融合方法。经过实验证明，该方法实用性好，数据平稳性好，外推稳定度高，已应用于卫星导航定位数据处理领域。

1 广义延拓插值外推法

广义融合是基于广义延拓插值外推法提出的，因此，先介绍广义延拓插值外推法。广义延拓插值外推法的基本设计理念是[9]：分段实施最小二乘逼近，在每段的端点处采用插值处理，在段内采用

* 基金项目：国家自然科学基金（61271284）；广西自然科学基金（2014GXNSFAA118393）；桂林电子科技大学研究生教育创新计划项目（GDYCSZ201483）资助.

收稿日期：2015-08-24；修订日期：2015-09-25

作者简介：魏彦飞，男，硕士. 研究方向：数据处理、导航定位. Email：729680894@qq.com

拟合处理，段内拟合逼近处理时，用邻近延拓域里的数据进行辅助拟合处理，从而使逼近多项式在定义域内的逼近精度升高；当分段连接后，在连接端点处过渡平滑，在整域上逼近精度比较好。广义延拓最小二乘逼近方法的外推模型兼融了插值和拟合两大功能，具有极好的逼近精度。外推示意图见图 1。

假设已知历元 t_n 时刻的最优估计值 $\hat{x}_n$ 和 t_n 之前的数据值($\hat{x}_1$，$\hat{x}_2$，…，$\hat{x}_{n-1}$)，求下一时刻 t_{n+1} 的值 x_{n+1}，这是一个典型的数据外推问题，可建立如下广义延拓插值外推模型[1]：

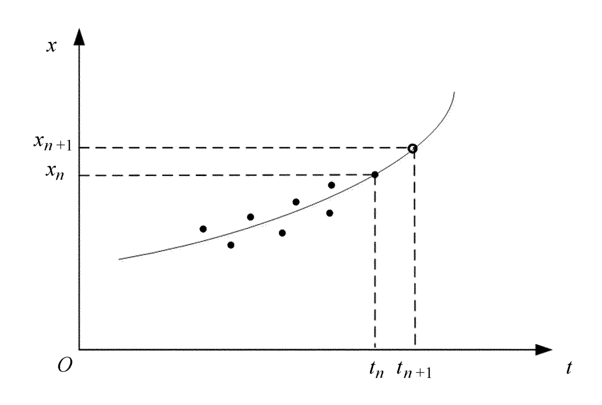


图 1　广义延拓外推示意图

Fig. 1　Generalized extrapolation diagram

$$\begin{cases} \min I(a_1, a_2, a_3) = \sum_{i=m}^{n-1} [a_1 + a_2 t_i + a_3 t_i^2 - x_i]^2 \\ s.t. \quad a_1 + a_2 t_n + a_3 t_n^2 = x_n \end{cases}, \tag{1}$$

其中，$\min I(a_1、a_2、a_3)$为逼近残差极小化目标函数；a_1、a_2、a_3 为待求的系数；$s.t.$示意约束条件。求解模型(1)，可以优化获得广义延拓逼近多项式的待求系数 a_1、a_2、a_3；代入(2)式：

$$x(t_j) = a_1 + a_2 t_i + a_3 t_i^2 \quad (i = n+1, n+2, \cdots), \tag{2}$$

便可以外推求解得到 t_{n+1} 历元的估计值 $\hat{x}_{n+1}$。

2　广义融合方法的数学模型及其求解方法

2.1　广义融合方法数学模型的建立

广义融合求解模型借鉴了广义延拓插值外推公式，即在 t_n 历元时，选用多项式 $a_1+a_2t_n+a_3t_n^2$ 作为逼近函数。逼近时选用一段历元内的数据，其中把已获得的先验最优估计值(起始时由于没有先验最优估计值，可以先用测量值代替)作拟合逼近处理；选用当前待求历元的广义状态量的预估值与测量值的最优加权组合逼近形式作为插值约束锁定点；以逼近多项式系数 a_1、a_2、a_3 和权系数 k_1、k_2 等为优化求解变量；以逼近函数多项式系数 a_1、a_2、a_3，权系数 k_1、k_2 以及广义状态量等待求量的区间数作为变量取值约束条件。这些约束条件把解限制在一定的范围之内，从而提高解的精度。

构造广义融合最优化求解模型，求 a_1、a_2、a_3，使满足：

$$\begin{cases} \min I_1(a_1, a_2, a_3) = \sum_{i=n-m}^{n-2} [a_1 + a_2 t_i + a_3 t_i^2 - \hat{x}_i]^2 \\ \qquad + k_1 [a_1 + a_2 t_n + a_3 t_n^2 - \tilde{x}_n]^2 + k_2 [a_1 + a_2 t_n + a_3 t_n^2 - x_n]^2 \\ s.t. \quad a_1 + a_2 t_{n-1} + a_3 t_{n-1}^2 = \hat{x}_{n-1};\ k_1 + k_2 = 1 \\ a_1 \subseteq \widehat{a}_1,\ a_2 \subseteq \widehat{a}_2,\ a_3 \subseteq \widehat{a}_3 \end{cases}, \tag{3}$$

其中，$I_1(a_1, a_2, a_3)$为待优化的目标函数，优化的目标函数既可以是线性关系也可以是非线性关系，若是非线性关系，则广义融合方法直接采用非线性求解算法求解；x_i 为状态量的测量值；$\hat{x}_i$ 为状态量的最优估计值；$\tilde{x}_i$ 为由低阶状态量递推获得的状态量的预估值；i 为历元序号；k_1、k_2 为权系数，它们的取值大小影响它们各自的残差对目标函数的贡献，权值的分配在本文后面作详细的介绍；x_n 为 t_n 历元时状态量的测量值；$\tilde{x}_n$ 为 t_n 历元时状态量的递推预估值；$\hat{x}_{n-1}$ 为 t_{n-1} 历元时状态量的最优估计值；$\widehat{a}_1$、$\widehat{a}_2$、$\widehat{a}_3$ 为可行区间变量 a_1、a_2、a_3 的约束区间数，$\widehat{a}_1 = [\underline{a}_1、\overline{a}_1]$，$\widehat{a}_2 = [\underline{a}_2、\overline{a}_2]$，$\widehat{a}_3 = [\underline{a}_3、\overline{a}_3]$。广义融合模型中可以包含非线性状态方程、非线性约束方程以及其它非

线性测量方程，这能够扩大应用范围，增加方程的冗余性和连续性，提高求解精度。

2.2 广义融合方法的求解步骤

2.2.1 求解状态量值

建立残差极小化优化模型并加约束方程，如下式：

$$\begin{aligned} &\min I(x_n) = \sum_{i=1}^{n} (F(x_i) - Y_i)^2 \\ &s.t. \quad f(x_n) = 0 \end{aligned}, \tag{4}$$

其中，$F(x_i)$为状态量关系式，多数情况下为非线性函数关系，可以采用非线性算法直接求解；x_i 为函数状态变量；Y_i 为测量量；i 为历元序号；n 为当前历元数；$f(x_n)$为约束函数关系式。(4)式可以采用非线性算法求解，如单纯形法等。通过搜索、比较等步骤，使目标函数值不断逼近最优点。求解方法在优化过程中不受方程个数的限制，并且具有简单直观、适应性强、应用范围广、解决速度快等特点。

2.2.2 求解状态量的预估值 $\tilde{x}_n$

利用外推得到的状态方程的高阶状态量可以提高解的精度和相关性，低阶状态量可以是与状态量相关的导数值或微分值。当求解 t_n 历元时刻状态量的预估值 $\tilde{x}_n$ 时，由于已经得到 t_{n-1} 历元时刻状态量的最优估计值 $\hat{x}_{n-1}$(在初始阶段的融合，初始状态最优估计值是未知的，可用前几个时刻状态的测量值代替)，可以通过公式：

$$\tilde{x}_n = \hat{x}_{n-1} + \dot{x}_n \Delta t + \frac{1}{2}\ddot{x}_n \Delta t^2 + \cdots , \tag{5}$$

递推获得 t_n 历元时刻状态量的预估计值 $\tilde{x}_n$。(5)式中 $\tilde{x}_n$ 为 t_n 历元时刻状态量的预估值；$\hat{x}_{n-1}$为 t_{n-1}时刻状态量的最优估计值；$\dot{x}_n$ 为 t_n 历元时刻状态量的一阶导数；Δt 为相邻历元之间的时间间隔；$\ddot{x}_n$ 为 t_n 历元时刻状态量的二阶导数。

2.2.3 求解状态量的最优估计值 $\hat{x}_n$

经上述步骤，已经求得状态量值和状态量的预估值，t_n 历元时刻直接求解获得的状态量 x_n 和 t_n 历元时刻获得的状态量的预估值 $\tilde{x}_n$，现在需要把它们组合起来，求得 t_n 历元时刻状态量的最优估计值 $\hat{x}_n$，即

$$\hat{x}_n = a_1 + a_2 t_n + a_3 t_n^2 , \tag{6}$$

其中，a_1、a_2、a_3 可以由广义融合最优化求解模型((3)式)采用直接法等优化算法求解得到。将得到的 a_1、a_2、a_3 代入(6)式可得到 t_n 历元时刻状态量的最优估计值 $\hat{x}_n$。

2.2.4 求解一组优化的最优估计值解

对上述4个步骤不断进行迭代、递推，直到结束，就能够得到一组优化的最优估计值解 $\hat{x}_i$，$i=1, 2, 3 \cdots k_{\text{end}}$。

3 广义融合模型之中权值分析

在广义融合方法数学模型((3)式)中 $a_1 + a_2 t_n + a_3 t_n^2 - x_n$ 表示逼近函数与状态量之间的残差；$a_1 + a_2 t_n + a_3 t_n^2 - \tilde{x}_n$ 表示逼近函数与状态量的预估值之间的残差。它们的权系数 k_1、k_2 的取值大小会影响它们各自的残差对目标函数的贡献，所以权值的选择极为重要。故在确定权系数之前要分析状态量的预估值 $\tilde{x}_n$、状态量的测量值 x_n 与逼近函数之间的残差大小比例关系；在融合数据时，要考虑被融合数据自身的精度及误差特性等诸多因素。通过改变 k_1、k_2 的大小可以调控广义融合最优估计值与状态量及预估值的靠近程度。

用仿真验证方法的可行性，即生成两组数据：一组数据随机误差较大但偏移较小(随机误差 -3-3，偏差为 1)；另一组随机误差较小但有较大的偏移(随机误差-1-1，偏差为 3)。数据生成后代

入广义融合求解模型，选取不同的4对权值进行组合分析，所得结果如图2。

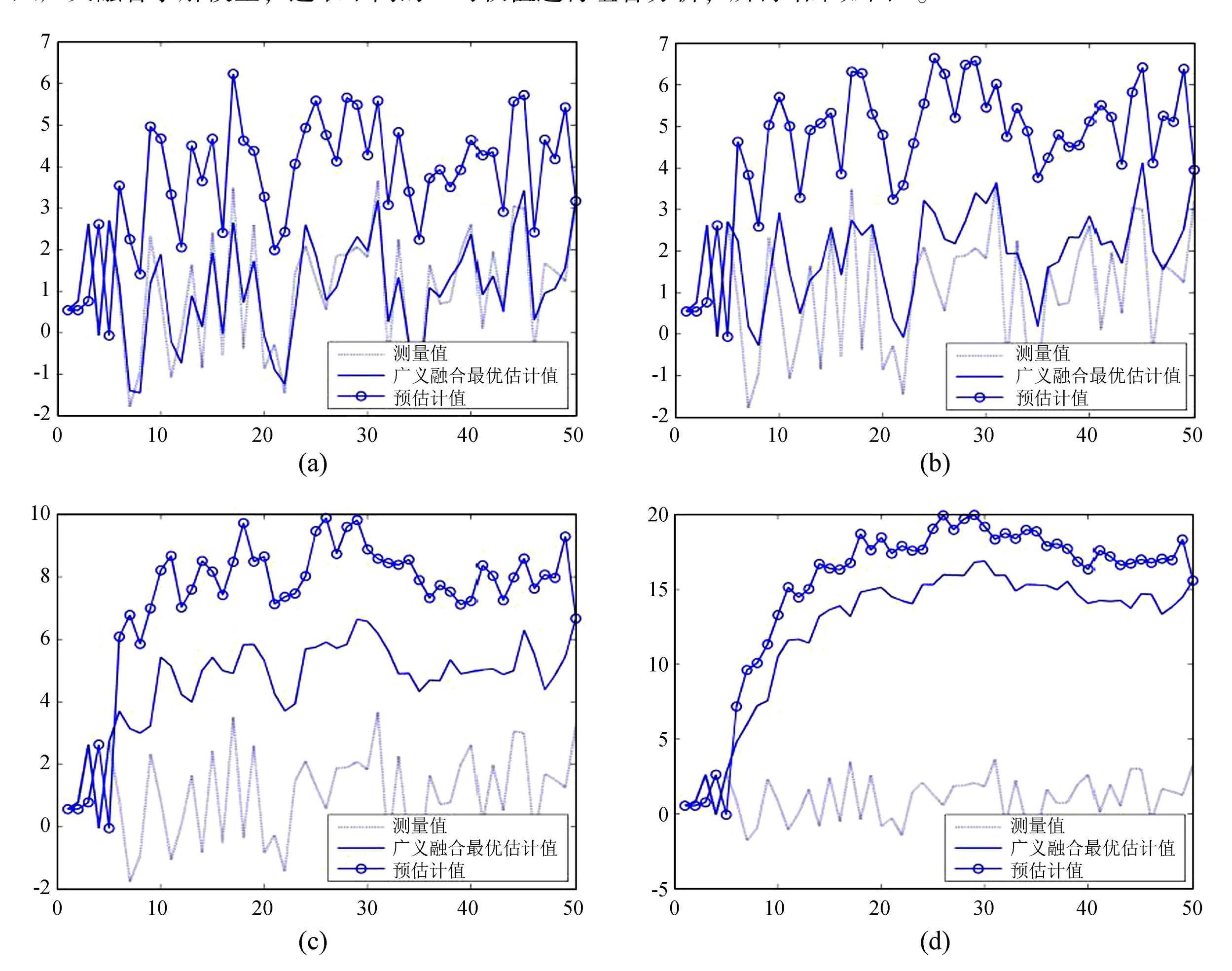


图2　(a) $k_1=1$、$k_2=0$ 时最优估计值偏移；(b) $k_1=0.7$、$k_2=0.3$ 时最优估计值偏移；(c) $k_1=0.3$、$k_2=0.7$ 时最优估计值偏移；(d) $k_1=0$、$k_2=1$ 时最优估计值偏移

Fig. 2　(a) $k_1=1$, $k_2=0$, offset of optimal estimation; (b) $k_1=0.7$, $k_2=0.3$, offset of optimal estimation; (c) $k_1=0.3$, $k_2=0.7$, offset of optimal estimation; (d) $k_1=0$, $k_2=1$, offset of optimal estimation

从图2(a)、(b)、(c)、(d) 4幅图可以看出：当 k_1 逐渐减小而 k_2 逐渐增大时，状态量的最优估计值开始远离状态测量值向预估值逼近；随着 k_1 越来越小，k_2 越来越大，这种现象更加明显。在数据处理过程中，如果有一组数据序列的某些测量量发生突跳导致数据不准确时，便可以通过改变权值组合来降低这些突跳数据对获取最优估计值的贡献，使获得的最优估计值更加平稳。

4　广义融合方法在卫星导航中的应用及仿真与实测分析

4.1　广义融合在卫星导航中的应用

以卫星定位导航为例说明采用广义融合方法的实用性：卫星导航定位解算出用户的位置坐标为 x、y、z 3个方向，本文只对一个方向的用户位置作详细分析，其余两个方向使用相同方法修正即可获得用户位置的完整坐标。

4.1.1　获得 t_m 时刻位置最优估计值解 $\hat{x}_m$ 和相应时刻的速度 $\hat{\nu}_m$

将从初始时刻到 t_m 历元时刻接收机的位置量设为 $x_i(i=1, 2, \cdots, m)$。通过对 t_m 历元时刻以前的位置测量量采用分段最小二乘拟合的方法(5个历元时刻点做一次拟合)对位置测量量进行处理，用处理后通过逼近多项式(6)获得的 t_m 时刻的值，作为 t_m 时刻位置的最优估计值解 $\hat{x}_m$[10]。由于在导航定

位中速度的测量值精确度较高，初始时刻的最优速度值及之后各时刻的最优速度值都直接选用对应时刻的速度测量值即可。

4.1.2 获取 t_{m+1} 时刻位置量的预估值 $\tilde{x}_{m+1}$

由第 1 步已经得到了 t_m 时刻位置的最优估计值解 $\hat{x}_m$ 和相应时刻速度的最优值 $\hat{\nu}_m$，利用(5)式解算出 t_{m+1} 时刻位置量的预估值 $\tilde{x}_{m+1}$。(5)式中 $\dot{x}_n$ 为 t_n 历元时刻状态量的一阶导数即 t_n 时刻的速度值；$\ddot{x}_n$ 为 t_n 历元时刻状态量的二阶导数即加速度值，加速度本文暂不考虑。

4.1.3 获得 t_{m+1} 时刻最优估计值解 $\hat{x}_{m+1}$

通过上述步骤，已经得到 t_{m+1} 时刻位置量的预估值，相应时刻位置量的测量值可以通过测量数据直接获得。因为实际的测量值会因为误差等因素导致最终结果不精确，所以本文在测量值与预估值之间添加一个阈值。当两者之间的绝对差值超过这个阈值时，便认为测量值已经发生了跳变，此时(3)式中的权值组合 $k_1 \approx 0$、$k_2 \approx 1$，使最优估计值的解向预估值方向逼近；当两者之间绝对差值没有超过阈值时，认为测量数据在误差范围之内，就要选择合适的权值组合，再代入(3)式，组合求解获得最优估计值 $\hat{x}_{m+1}$。

4.1.4 获得经优化的一组最优估计值解

上述第 2、第 3 步不断迭代、递推、融合就能获得一组优化的最优估计值解 $\hat{x}_{m+1}$，$\hat{x}_{m+2}$，…，$\hat{x}_{\text{end}}$。

4.2 仿真

4.2.1 广义融合对突跳数据的修正

卫星定位系统的精度主要取决于两方面：一是观测卫星在空间的分布情况，通常称为卫星分布几何衰减因子；二是各观测量的精度[11]。大气层干扰、对流层延迟改正后的残差、星历误差及多路径效应等影响导致最后的测量数据产生偏差或是突跳，而广义融合方法能够很好地修正这些误差数据。现生成一组随机误差在(－0.5-0.5)的数据，分别在第 10、20、30、40 4 个点处加上较大的偏移量构成突跳点，然后用广义融合方法对这组数据进行处理，仿真结果如图 3。

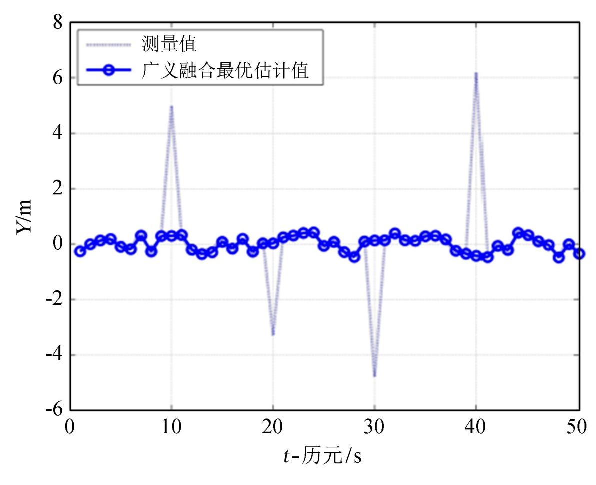


图 3 广义融合对突跳数据的修正

Fig. 3 Correction of the data of sudden jump with generalized fusion

由图 3 可以看出，广义融合法对 4 个突跳点的数据都做了实时修正，其结果不但在没有发生突跳的部分很好地遵循了原始测量值的变化规律，而在有突跳的时段，也能有效排除突跳的影响。

4.2.2 最小二乘拟合法与广义融合法对误差数据处理对比

在卫星数目多于 4 颗，各测量值之间相互独立且符合高斯正态分布的情况下，采用最小二乘算法解算接收机的三维坐标是最经典的办法。因为最小二乘算法能使各观测伪距之间的残差平方和最小，从而使获得的估计值解达到最优。但是多路径效应等因素产生的误差却时刻影响其定位精度，这导致各实际测量值之间并不是相互独立，甚至有些情况下不符合正态分布，这时若仍采用最小二乘算法就很难保证最后能获得最优估计值解[12]，而广义融合法在此种情况下依然适用。

下面通过仿真实验验证此方法的有效性。令测量值的理论值为 50，这时在理论值的基础上加噪声生成一组带有随机误差的数据，作为测量仿真计算数据，为了便于体现广义融合法的优越性，分别用广义融合法和最小二乘算法对这组数据进行处理，修正后的效果如图 4。

经计算，得到最小二乘拟合值的标准差为 $\sigma = 3.626\,396\,349\,003\,966$，广义融合最优估计值的标准差为 $\sigma = 0.758\,068\,193\,045\,756$，广义融合法稳定程度要比最小二乘拟合法提高大约 79%(图 4)；而且广义融合法的偏差也较小，误差数据的修正值更加靠近理论值。

4.2.3　实测数据分析

(1)2015 年 5 月 11 日在中国科学院国家天文台用中科微 ATGM332D 卫星定位芯片的接收机做动态测试。从最终获取的接收机三维坐标中截取 Y 方向上一段带有明显突跳的原始定位数据和对应时刻的原始速度值，分别用广义融合法和最小二乘拟合法对其进行处理，结果如图 5。

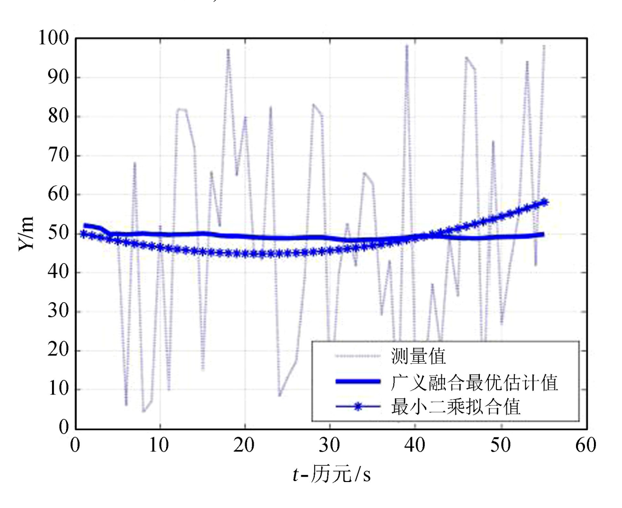


图 4　最小二乘拟合法与广义融合法对误差数据处理的对比

Fig. 4　Comparison of error data processed by the least squares fitting method and the generalized fusion method

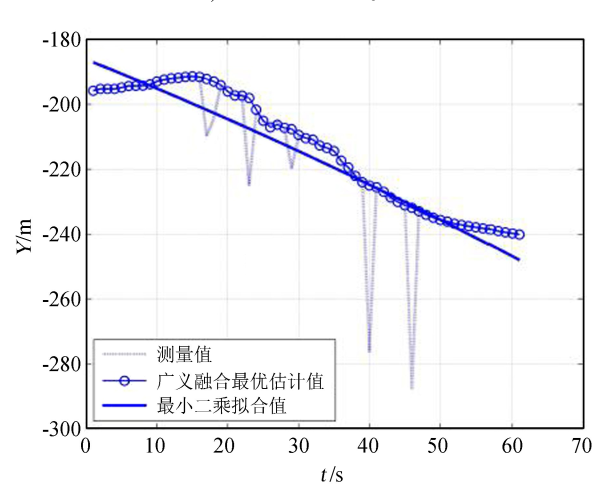
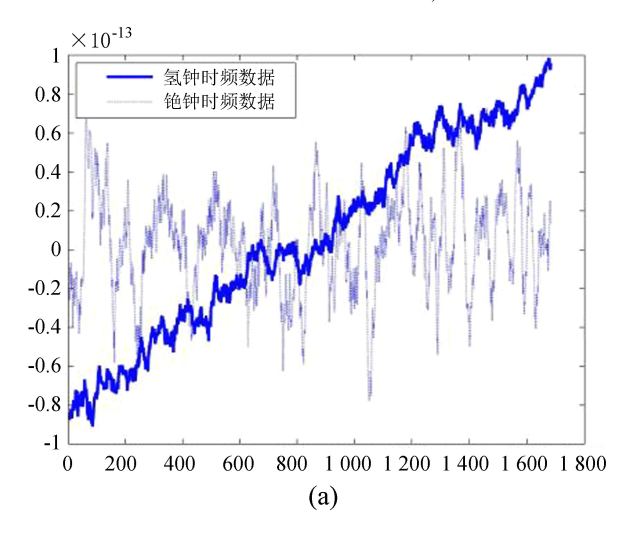


图 5　广义融合法和最小二乘拟合法对实测数据处理的对比

Fig. 5　Comparison of measured data processed by both the generalized fusion method and the least square fitting method

由图 5 可以看出最小二乘拟合法对某些突跳点并没有很好的修复作用，只是在少数点处遵循了原始数据的变化规律，但是整体数据偏离了实测数据；而广义融合法能实时修正原始数据中的坏数据，并且在没有发生突跳的部分很好地跟随原始数据的运动轨迹。

(2)当代社会对时频基准的要求越来越高，氢原子钟、铯原子钟等多种高精度时频基准已被相继研发出来。然而还可以通过另外一种方法获得更高精度和稳定度的时频基准，那就是组合处理和精细加工原子钟组的输出数据，获得性能更优良的时频输出数据。本文把氢原子钟与铯原子钟的输出数据用广义融合法融合起来。图 6(a)显示了铯原子钟与氢原子钟输出的频率数据变化图，通过对比可以发现铯原子钟频率的随机误差偏大，但长期稳定度很好；而氢原子钟则不同，它随机误差的幅度很小，



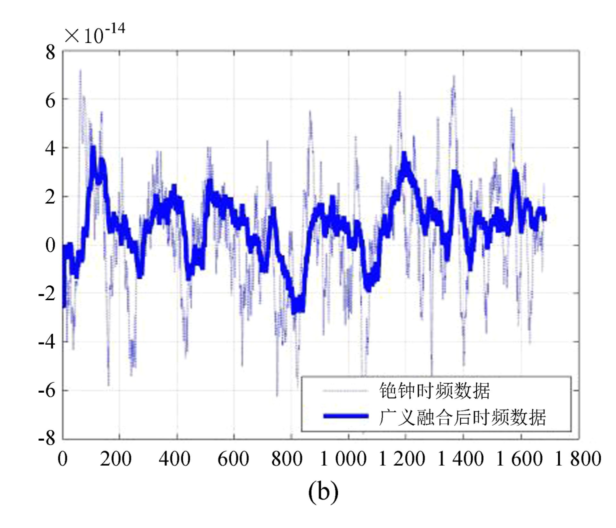


图 6 (a) 氢钟、铯钟时频数据对比；(b) 广义融合前后时频数据对比

Fig. 6　(a) Frequency data of hydrogen and cesium clock; (b) Comparison of time frequency before and after the generalized fusion

但是随着时间的增长其频率出现很明显的漂移。广义融合法能把这两种原子钟的频率特性的优势组合起来，从而获得随机误差小并且长期稳定度好的时频基准。用这一方法对国家授时中心的铯原子钟和氢原子钟输出的时频数据进行融合处理，用氢钟时频的变化率修正铯钟的时频输出，处理结果见图 6(b)。

5 结 论

本文针对卫星导航定位数据中出现的坏数据，尤其是突跳数据的修正，在广义延拓插值外推法的基础上提出了一种新的数据融合方法——广义融合法。通过仿真和实测数据分析，与最小二乘算法进行对比后得出：广义融合法较最小二乘拟合法更能够修复定位数据中产生的突跳数据；广义融合法比最小二乘拟合法对数据的处理可靠性更高；广义融合法实用性好，数据平稳性好，外推稳定度高，有望成为一种新的解决数据融合问题的普适求解方法。

参考文献：

[1] 王栋. 基于数据融合的机载多传感器目标识别 [D]. 上海：上海交通大学，2010.

[2] 王炯琦，周海银，吴翊. 基于最优估计的数据融合理论 [J]. 应用数学，2007，20(2)：392-399.
Wang Jiongqi，Zhou Haiyin，Wu Yi. The theory of data fusion based on state optimal estimation [J]. Mathematica Applicata，2007，20(2)：392-399.

[3] 石忠，霍星. 多传感器信息融合技术在 INS/GPS 组合导航系统中的应用 [J]. 渤海大学学报：自然科学版，2008，29(4)：403-407.
Shi Zhong，Huo Xing. Application of multi-sensor information fusion technique in INS/GPS integrated navigation system [J]. Journal of Bohai University：Natural Science Edition，2008，29(4)：403-407.

[4] Shang Lin，Liu Guohua，Zhang Rui，et al. An information fusion algorithm for integrated autonomous orbit determination of navigation satellites [J]. Acta Astronautica，2013，85：33-40.

[5] Ning Xiaolin，Fang Jiancheng. An autonomous celestial navigation method for LEO satellite based on unscented Kalman filter and information fusion [J]. Aerospace Science and Technology，2007，11(2-3)：222-228.

[6] Tzschichholz T，Boge T，Schilling K. Relative pose estimation of satellites using PMD-/CCD-sensor data fusion [J]. Acta Astronautica，2015，109：25-33.

[7] Wang Xinlong，Zhang Qing，Li Hengnian. An autonomous navigation scheme based on starlight，geomagnetic and gyros with information fusion for small satellites [J]. Acta Astronautica，2014，94(2)：708-717.

[8] 施浒立，颜毅华，徐国华. 工程科学中的广义延拓逼近法 [M]. 北京：科学出版社，2005.

[9] 耿建平，衣伟，刘成，等. 基于广义延拓外推的单频周跳检测与修复方法 [J]. 天文研究与技术，2015，12(2)：174-182.
Geng Jianping，Yi Wei，Liu Cheng，et al. A new method for detection and correction of single-frequency cycle slipsusing data extrapolationbased on the method of generalized extended interpolation [J]. Astronomical Research & Technology，2015，12(2)：174-182.

[10] 施浒立，黄康，衣伟，等. 一种新的逼近滤波方法：中国，201410208791.7 [P]. 2014-08-06.

[11] 何瑞珠，刘成，黄康．卫星定位中的一种分步加权解算方法［J］．天文研究与技术，2015，12(1)：36-43.
He Ruizhu，Liu Cheng，Huang Kang. A method of multiple-step Weighted Least-Square estimate for satellite positioning［J］. Astronomical Research & Technology，2015，12(1)：36-43.

[12] 朱丽梅，张绪春．高动态卫星接收机定位高程超差问题分析［J］．海军航空工程学院学报，2010，25(2)：203-207.
Zhu Limei，Zhang Xuchun. Analysis of the positioning error on elevation during the high dynamic envenoming［J］. Journal of Naval Aeronautical and Astronautical University，2010，25(2)：203-207.

A New Method of Data Fusion——the Generalized Fusion Method

Wei Yanfei[1,2]，Geng Jianping[1]，Shi Huli[2,3]，Li Jingjing[4]，Li Lin[2]，Xu Xikun[1]

(1. Electronic Engineering and Automation College，Guilin University of Electronic Technology，Guilin 541004，China；
2. National Astronomical Observatories，Chinese Academy of Sciences，Beijing 100012，China，Email：729680894@qq.com；
3. Beijing Ri Yue Jiu Tian Technology Co.，Ltd，Beijing 100012，China；4. Electronic Information College，
Zhongyuan University of Technology，Zhengzhou 450007，China)

Abstract：The purpose of data fusion is to merge or synthesize two data sets so as to make them become a better and complete data set. Based on extended inter-polation method，generalized fusion method is a method which strives to create a mathematical model that consists of two or more sets of equation and data sets. This method is not only able to bring out its full potential of data and measurement resources，but also capable of using innovative integration and mutual algorithm to solve the data fasion problem. After simulation and real measurement data analysis，conclusion can be drawn that compared with the least square algorithm，the generalized fusion method is more suitable for the repair of the data generated in the process of data fusion. This paper also introduces the main steps of the establishment and the solution of the generalized mathematical model，analyzes the distribution principle of weight in the model and its application in the field of satellite navigation data processing. Experimental results show that the method is practical and stable，and it is expected to be a new method for solving the problem of data fusion.

Key words：Data fusion；Generalized fusion；Generalized extrapolation；Model solving；Weight distribution；Satellite navigation data processing